(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 298 769 A2

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 88306270.5

(22) Date of filing: 08.07.88

(51) Int. Cl.⁴: B 07 C 1/00

(30) Priority: 08.07.87 JP 170198/87

(43) Date of publication of application:
11.01.89 Bulletin 89/02

(84) Designated Contracting States: FR GB NL

(71) Applicant: KABUSHIKI KAISHA TOSHIBA
72, Horikawa-cho Saiwai-ku
Kawasaki-shi Kanagawa-ken 210 (JP)

(72) Inventor: Ota, Naoki c/o Patent Division
Toshiba Corp. Principal Office
1-1, Shibaura 1-chome Minato-ku Tokyo (JP)

(74) Representative BATCHELLOR, KIRK & EYLES
2 Pear Tree Court Farringdon Road
London EC1R 0DS (GB)

(54) Mail processing machine.

(57) A machine for processing mail to separate the items whose destination address is typed or printed from those which are in manuscript. The machine reads the information on the items and distingutshes the destination address from a sender's address. The destination address is then examined to recognise typed or manuscript forms according to the dispersion of the height of letters, the spacing between letters or lines, etc. The machine calculates the dispersion of only the bottom or one or two lines of the address because this is where the most useful information is present.

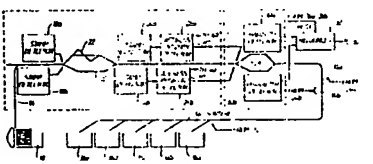

Fig 2

EP 0 298 769 A2

Description

MAIL PROCESSING MACHINE

This invention relates to aparatus for, and a method of, processing mail to select from a multiplicity of items of mail to be processed those items having a destination address comprising a plurality of characters which are printed or typed thereon.

Mail processing apparatus employs means for recognising a destination address on each item of mail. Where the destination address is printed, and that includes typing, on the item of mail, it can readily be recognised but, where the destination address is written in manuscript form on the items of mail, the recognition apparatus has difficulty in recognising the destination address. In prior art mail processing apparatus, the items of mail have to be sorted manually in order to separate those items of mail which have a printed destination address from those having an address in manuscript. It will be appreciated that this is a time consuming operation.

It is an object of the present invention to provide apparatus and a method for selecting from a multiplicity of items of mail those having a printed or typed destination address, thus separating those items of mail from items having a destination address applied in manuscript.

According to a first aspect of the present invention, mail processing apparatus comprises

means for detecting on each item of mail the position of an area which contains the destination address;

means for generating signals representative of characters within the destination address area;

means for calculating the dispersion of a characteristic of the characters within the destination address area;

means for comparing the dispersion with a predetermined level; and

means for selecting those items of mail as having a printed or typed address where the dispersion is less than the predetermined level.

According to a second aspect of the invention, a method of processing mail comprises the steps of

(a) detecting on each item of mail the position of an area which contains the destination address;

(b) generating signals representative of characters within the destination address area;

(c) calculating the dispersion of a characteristic of the characters within the destination address area;

(d) comparing the dispersion with a predetermined level; and

(e) selecting those items of mail as having a printed or typed address where the dispersion is less than the predetermined level.

In order that the invention may be more readily understood, it will now be described, by way of example only, with reference to the accompanying drawings, in which:-

Figure 1 is a plan view of an item of mail which carries a printed address of the sender of the mail and a hand written address for the recipient of the item of mail;

Figure 2 shows diagrammatically an embodiment of the mail processing machine according to the present invention;

Figure 3 is a block diagram of the machine shown in Figure 2;

Figure 4 illustrates diagrammatically four alternative positions of items of mail fed into the machine shown in Figure 2;

Figure 5 shows diagrammatically an address position detector shown in Figure 2;

Figure 6 is a graphical representation of signal levels detected by the address position detector shown in Figure 5;

Figure 7 is a block diagram showing a character detector and recognising apparatus shown in Figure 2;

Figures 8 to 10 show diagrammatically how the recoganiser of Figure 2 separates a line or character;

Figures 11 to 16 are illustrations explaining the calculation of the dispersion of a characteristic of characters of the recogniser shown in Figure 2; and

Figure 17 is a block diagram showing address position detector of Figure 7.

Referring to Figure 1, an item of mail P is in the form of a generally rectangular envelope. On the front of the envelope in the top right-hand corner there is a postage stamp. On the front left-hand side of the envelope there is a name and address printed on the envelope and on the right-hand side of the envelope there is name and address applied in manuscript to the envelope. It is assumed that, because the stamp is positioned in the top right-hand corner, the manuscript name and address is that of the destination of the item of mail, whereas the printed name and address is that of the sender.

Apparatus for processing items of mail is shown in Figure 2 and the control system is shown in more detail in Figure 3.

A multiplicity of items of mail are arranged side-by-side on edge in a mail box 38. Each item is fed, in turn, from the mail box along a path 1a to a position between a pair of stamp detectors 18a and 18b. The items of mail are arranged vertical and, in this position, the postage stamp, which is normally on the front top right-hand corner of the envelop, can be in any of the positions a, b, or $\bar{a}$, or $\bar{b}$, as shown in Figure 4. In other words, the item can be arranged with the side of the envelope which carries the stamp facing either detector 18a or detector 18b and, on that side, the stamp may be in the bottom left-hand corner or in the top right-hand corner. If the stamp is detected by either of the detectors 18a or 18b, it means that the stamp is positioned at or towards the bottom of the item of mail and it is passed along the path 20. If, however, neither of the detectors detect the presence of a stamp, it means that the stamp is at the top side of the item and the

item is passed along an inversion path 22 where the envelope is inverted so that the stamp is now towards the bottom of the envelope. The items of mail then pass between a pair of stamp detectors 24a and 24b to make sure that, in fact, a stamp is present on the envelope.

Address position detectors 26a and 26b receive signals from the detectors 24a and 24b, respectively, and the detectors 26a and 26b detect the presence of either an address window covered by cellophane or an address label which indicates the presence of an address on the envelope but, where the presence of a stamp is not detected, the quantity of the characters or the extent of the characters written on one surface of the item of mail is compared with that on the opposite surface of the item in order to determine the front surface or the back surface of the envelope. That is, that surface on which the most characters are detected is considered to be the front surface of the envelope.

When the detector 26a detects the front of the envelope, a branch mechanism 28 is operated so that the envelope is diverted along a path to the character detector 30a. On the other hand, when the address position detector 26b detects the front of the envelope, the branch mechanism 28 is operated so that the envelope is fed to the character detector 30b.

Referring now to Figure 5, the address position detectors 26a, 26b comprise a light source 262 which emits a light beam towards the items of mail fed along a path, a lens 264 for focusing light scattered by the items of mail, a photosensitive element 266 composed of line image sensor (e.g. charged coupled devices) for detecting characters on the mail, an amplifier 268 and two comparison circuits 270a and 270b. The comparision circuit 270a compares the image signal S detected by the element 266 with a reference signal B. If the signal S is greater than the signal B, the comparison circuit outputs a high level signal. Such an arrangement occurs when the signal Sc shown in Figure 6 exceeds the level B. This is likely to occur when there is a window or label of high reflectivity on the envelope. On the other hand, the comparison circuit 270b compares the image signal S with a signal C at a lower reference level than B and a signal Sb is indicative of the absence of characters of medium reflectivity on the envelope.

Further, when the level of the image signal S lies between the reference level C and a lower level A, the comparison circuit 270b outputs a high level signal which is called a character signal indicative of the presence of characters of low reflectivity.

In the address position detector 26a or 26b, the resolving power of scanning is not high (e.g. a single scanning line per mm) because this detector detects only the position of an address window or a label and the window or label can be detected in the form of x y coordinates indicative of the number of the horizontal scanning lines from an upper edge and a time period from an edge of the horizontal scanning line.

When no postage stamp and no address window or label is detected, the character signal Sd outputted from the comparison circuit 270b is integrated by an integrator 272 and supplied to a comparator 274. A paper surface signal Sb is not integrated by the integrator 272 but a character signal Sd outputted from the comparison circuit 270b of the address position detector is integrated by the integrator 272 and supplied to the comparator 274. The comparator 274 compares the two integrated numbers of character signals to determine the front side of an envelope where it is supposed that the address is written. For instance, if the number of character signals integrated by the integrator 272 is larger than that by the integrator 272', the comparator 274 generates a command signal to the branch mechanism 28 to feed the mail toward the character detector 30a. In response to this command signal, character detector 30a is activated to detect the character image.

Fig 7 shows the two character detectors 30a and 30b and the recognizer 32. Each character detector 30a or 30b comprises a fine scanner 302a or 302b and a quantization circuit 304a or 304b. The fine scanner 302a or 302b generates image signals in almost the same way as in the address position detector 26a or 26b by irradiating the mail surface with a light beam and transducing the reflected light beam (not scattered light beam) by photosensitive elements into image signals. However, the resolving power of the fine scanner 302a or 302b is as high as 8 lines per milimeter because this detector detects the features of characters.

The quantization circuit 304a or 304b compares the detected mail surface image signals with a predetermined reference level and outputs character image signals only when the image signal drops below a reference level (the aboe processing being referred to as binarization).

The recognizer 32 comprises an image memory 322, an address position detector 324, a line detector 326, a parameter extractor 328, and a discriminator 330.

The image memory 322 stores all the scanned and binatized character image signals detected by either one of the chatacter detector 30a or 30b. This is because the front surface of a mail has already been detected by the address position detectors 26a and 26b, and the detected mail is fed to any one of the character detectors 30a and 30b. Therefore, the image memory 322 stores the character image signals corresponding to the detected front surface of a mail and detected by any one of the character detectors 30a and 30b.

The line detector 326 functions as follows: The preceeding processings have already detected an address position or area where an address is written. Therefore, in this step, character lines are further detected from the detected address area. That is, since an area where characters are gathered has alredy been determined, the succeeding step determines how the cahracters are arranged within the detected address area.

For doing this, the number of character image signals are counted along the direction in order to obtain a histogram as shown in Fig. 8. By detecting the minimum of the histogram indicative of the

distribution of the character imge signals, it is possible to separate each of lines so that the position of coordinate is gotten.

The parameter extractor 328 detects character feature parameters. These parameters are dispersions of various character features such as (1) character height; (2) character lower edge position; (3) character width; (4) character pitch; (7) leftmost character position; (8) line space etc.

To obtain character feature parameters, each area of the character is separated within a line in an address area. In more detailed, with reference to Fig. 9, the number of image signals are counted along the direction X perpendicular to a line. By detecting the minimum of the histogram indicative of the distribution of the image signals, it is possible to separate each of characters so that the position number of the cordinate is gotten. It is possible to separate each area where each character is within each line with the position number of the cordinate and abscicaa as shown in Fig 10.

For example, as shown in Fig. 11 by the address position detector 26a, 26b, an address area 4D1 is specified as area with the upper leftmost position A and the low rightmost position (not shown). The separated character is specified as a character area with the upper left position and the low rightmost psition or the coordinate with the starting position A. The character area a is specified with (a_{x1}, y_1), (a_{x2}, y_2). The character area b is specified with (b_{x1}, y_1), (b_{x2}, y_2).

The dispersopns 2 of character features are calculated as follows:

$$\sigma_L^2 = 1/N \sum_1^n (X_{Li} - \overline{X_{Li}})^2 \quad (1)$$

N indicates the number of characters. L indicates the character feature such as height, lower edge position, width, pitch, area, line arrangement slope, leftmost position, line space and so on.

X_{Li} indicates the value of the character feaure of each character.

$\overline{X_{Li}}$ indicated the averaged value of the character feature of characters.

General speaking, σ_L^2 is larger in handwriting than that in printing. Fig. 12 shows an example presenting how the dispersion of character height σ_H^2 is in printing or handwriting. Fig 13 to Fig 16 show examples presenting how the dispersion of character lower edge position, character width, character pitch, character area are in printing or handwriting.

When calculating these feature parameters, all of the position number of characters are generally used.

But these calculating are a burden for a calculator (CPU) so that a calculating time is long and a calculator is needed to calculate at high speed.

In more desirable embodiment, the parameter extractor 328 may detect feature parameters on the basis of signals from a limited, within the address area. Especially, the restricted area is two lines from the bottom of the address area. There are imppportant information about the destination address in these two lines.

On the other hand, a reference threshold value σ_O^2 of each of the dispersions of the character features is previously determined. Each actual dispersion value σ_L^2 obtained by calculating image signals read from the image memory 322 is compared with this reference threshold value σ_O^2. The compared result (the difference between the actual dispersion and the reference dispersion) is stored in the image memory 322 and added in sequence to obtain a sum total of the differences between the two of the above-mentioned eight character features. When the discriminator 330 determines that the sum total of the dispersion differences between the actual values and the reference values exceeds a predetermined value, the characters are discriminated as a handwritten mail. In contrast with this, when the discriminator 330 determines that the sum total of the diapersion differences is less than the predetermined value, the characters are discriminated as a printed mail.

Fig. 17 is a block diagram showing the address position detector 324, which comprises a W/L signal detector 3242, a compressor 3244, an address area detector 3246.

When the W/L signal detector 3242 detects a presence of W/L signal indicative of a window/label position (x-y ccoordinates), the image data corresponding to only the window/label positon from the image memory 322 is supplied to the line detector 326.

When the W/L signal detector 3242 detects an absence of W/L signal, the compressor 3244 reads the entire image signals from the image memory 322 for compression. For instance, the resolving power of the image signals is reduced from 8 lines per millimeter to 1 line per millimeter by simply averaging the eight horizontal scanning line signal levels. The address area detector 3246 compares averaged signal levels with a reference level for binarization as same as the quantization circuit 270, and determines an address character area on the basis of the binalized character image signals collected at an area on the front surface of a mail. When this address character area has been detected, the image data corresponding to only the determined address character area is supplied to the line deterctor 326.

When the stamp detector 24a or 24b detects the presence of a postage stamp on the mail, a stamp canceller 34a or 34b corresponding to the stamp detecter 24a and 24b impresses a mark on the detected postage stamp. The mails thus detected are sorted and put into five sorting boxes 36a to 36e, in such a way that mails having an address written in print and detected by the character detector 30a are arranged in the box 36a; mails having an address written in handwriting and detected by the character detector 30a are arranged in the box 36b; mails having an address written in print and detected by the character detector 30b are arranged in the box 36c; mails having an address written in handwriting and detected by the character detector 30b are arranged in the box 36d; and other mails determined to be rejected are arranged in the box 36e.

In the machine of the present invention, it should be noted that since the address position detector 26a or 26b can detect a window/label position and the front side of the mail (by comparing the quantity

of characters) and further the address position detector 324 can determine an address character area, character images corresponding to only the front surface of the mail and only the address position (window or label) or address character area can be read from the image memory 322 for discrimination. Therefore, character image data to be discrimimated ara not huge, so that it is possible to increase the mail processing speed and therefore decrease the cost of the machine by providing only a single recognizer 32 including the discriminator 330.

The operation of the mail processing machine of the present invention for a mail as shown Fig. 1 will be described hereinbelow.

The mail P is arranged in the mail box 38 and fed one by one to the stamp detectors 18a and 18b via a path 1A in a vertically arranged position. When the stamp detector 18a or 18b detects the presence of a stamp attached to the lower side end of the mail, for instance, the mail is fed through the non-inversion path 20. When the stamp detector 18a or 18b detects the absence of a stamp, the mail is fed through the inversion path 22. Thereafter, the stamp detector 24a or 24b detects the presence or absence of a stamp is detected, this stamp presence signal is applied to the stamper 34a or 34b to impress a mark on the stamp of the mail just before sorting the mails.

When no stamp is detected by the two stamp detectors 24a and 24b, the front side of the mail (on which an address is written) is detected by the two address detectors 26a and 26b.

The address position (surface information) detector 26a or 26b also detects the position of a window or a label. In this process, when the address position detector 26a detects a mail front, the brance mechanism 58 is actuated so that the mail is fed to the character detector 30a; and when the address position detector 26b detects a mail front, the branch mechanism 58 is actuated so that the mail is fed to the character detector 30b.

Since the front surface of a mail has already been detected by the address position detectors 26a and 26b and the detected mail is fed to any one of the character detectors 30a and 30b. The character detector 30a or 30b detects characters on the front surface of a mail by scanning and quantization. The detected character image signals detected by the charecter detector 30a or 30b are stored in image memory 322. Further, only the character image signals corresponding to the address position signals (window/label signal) are supplied from the image memory 322 to the line detector 326 on the basis of the window/label signal detected by the address position detector 20a or 20b.

The character features (e.g. arrangement order, regularity, size, density, etc.) of the character images are detected by the parameter extractor 328 and discriminated as to printed mail or handwritten mail by comparing the extracted character futures with the stored reference character values by the discriminator 330.

For a mail P, where no window/label signal is detected, the address position detector 324 itself determines an address area by compressing the entire surface image signals and binarizing the compressed signals.

The process of separating each of lines is done for the signals of the destination address. So the signals are separated into the many group of signals corresponding to lines. As mentioned above the signals corresponding to a line are separated into a group of signals corresponding to a character. But in this embodiment, the signals but corresponding to two lines are used not corresponding to full lines in an address area. These two lines signals are used in descrimation.

After compressed, the signals exceed the threshold value to point out these existing a plural groups of signals. These groups of signals may come from destination address, a sender address an advertisement and so on. For the mail P, there are two areas of signals. By the knowledge of there existing a stamp on the right-upper side on the mail P in Fig.1, the right area is supposed to be a destination address. Because there is a destination address on the right side generally in the case that there is a stamp on the right-upper on the mail P.

The mails thus discriminated are stored into the five sorting boxes 36a to 36e.

In the above description, the mail processing machine of the present invention has been disclosed with reference to block diagrams (i.e. hardware configuration). In practice, however, the mail processing machine is controlled by the controller 12 provided with a ROM, a RAM, a display unit, a keyboard (not shown), etc., which is operated in accordance with control programs (i.e. software).

As described above, in the mail processing machine of the present invention, since the surface information (window or label position, quantity of characters, character block position) is first detected and then only the character image signals limited by the surface information are discriminated as to whether the address characters are written in printing or handwriting, it is possible to improve the sorting speed of the mail, while reducing the cost of the machine.

Claims

1. Mail processing apparatus for selecting from a multiplicity of items of mail to be processed those items having a destination address comprising a plurality of characters which are printed or typed thereon, said apparatus comprising

means for detecting on each item of mail the position of an area which contains the destination address;

means for generating signals representative of characters within the destination address area;

means for calcualting the dispersion of a characteristic of the characters within the destination address area;

means for comparing the dispersion with a

predetermined level; and

means for selecting those items of mail as having a printed or typed address where the dispersion is less than the predetermined level.

2. Apparatus as claimed in claim 1, wherein the calculating means comprises means for calculating the dispersion of a characteristic of the characters within only a portion of the destination address area.

3. Apparatus as claimed in claim 1 or 2, wherein the means for detecting the position of the destination address area comprises means for detecting the position of a high reflective portion of the item of mail and for specifying the high reflective portion as the destination address area.

4. Apparatus as claimed in claim 1, 2 or 3, wherein the means for calculating dispersion comprises means for calculating the dispersion of a characteristic of the characters, said characteristic being selected from character height, character lower edge portion, character width, character pitch, leftmost character position, and character line spacing.

5. A method of processing mail to select from a multiplicity of items of mail those items having a destination address comprising a plurality of characters which are printed or typed thereon comprising the steps of

(a) detecting on each item of mail the position of an area which contains the destination address;

(b) generating signals representative of characters within the destination address area;

(c) calculating the dispersion of a characteristic of the characters within the destination address area;

(d) comparing the dispersion with a predetermined level; and

(e) selecting those items of mail as having a printed or typed address where the dispersion is less than the predetermined level.

6. Method as claimed in claim 5, wherein the calculating step comprises the steps of calculating the dispersion of a characteristic of the characters within only a portion of the destination address area.

7. Method as claimed in claim 5 or 6, wherein the detecting step comprises detecting the position of a high reflective portion of the piece of mail and for specifying the high reflective portion as the destination address area.

8. Method as claimed in claim 5, 6, or 7, wherein the calculating step comprises the step of calculating the dispersion of a characteristic of the characters, said characteristics being selected from character height, character lower edge position, character width, character pitch, leftmost character position, and character line spacing.

9. Apparatus as claimed in claim 1 in combination with means for arranging the items of mail with the stamps arranged in a predetermined position and means for franking the stamps.

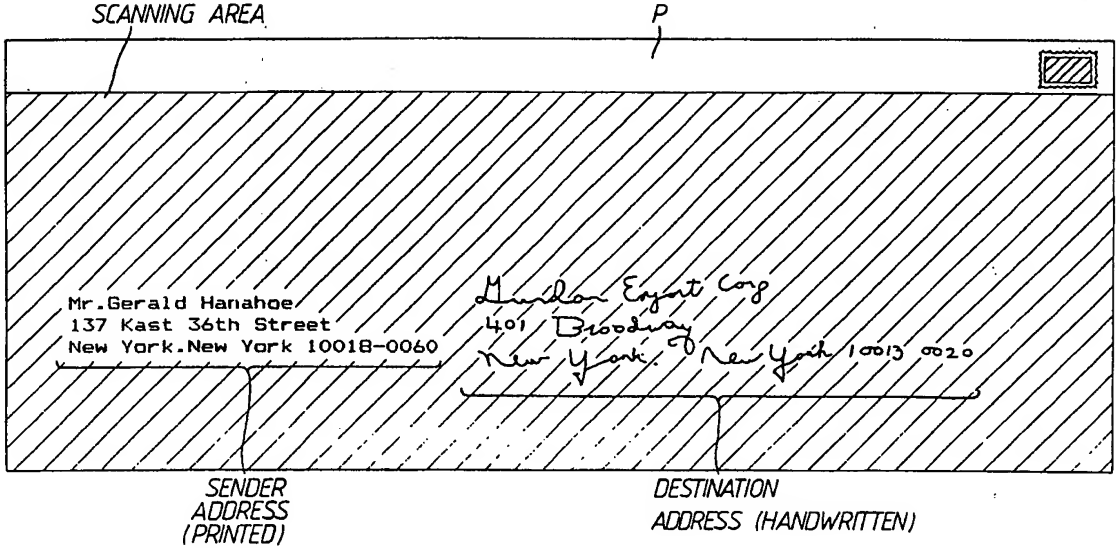

FIG.1.

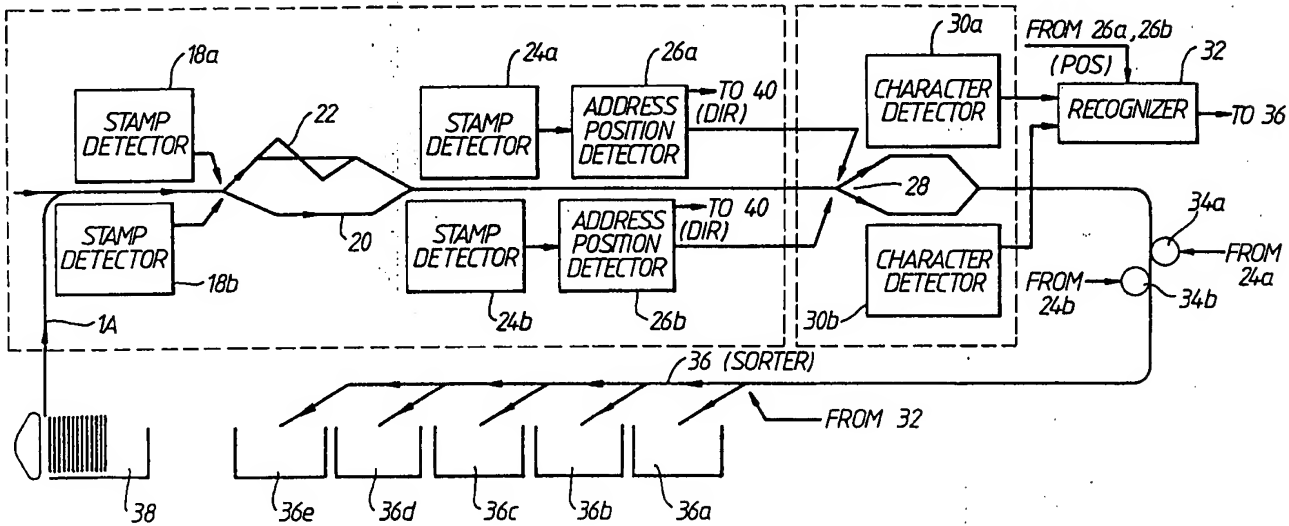

FIG.2.

0298769

24.10.88
0298769

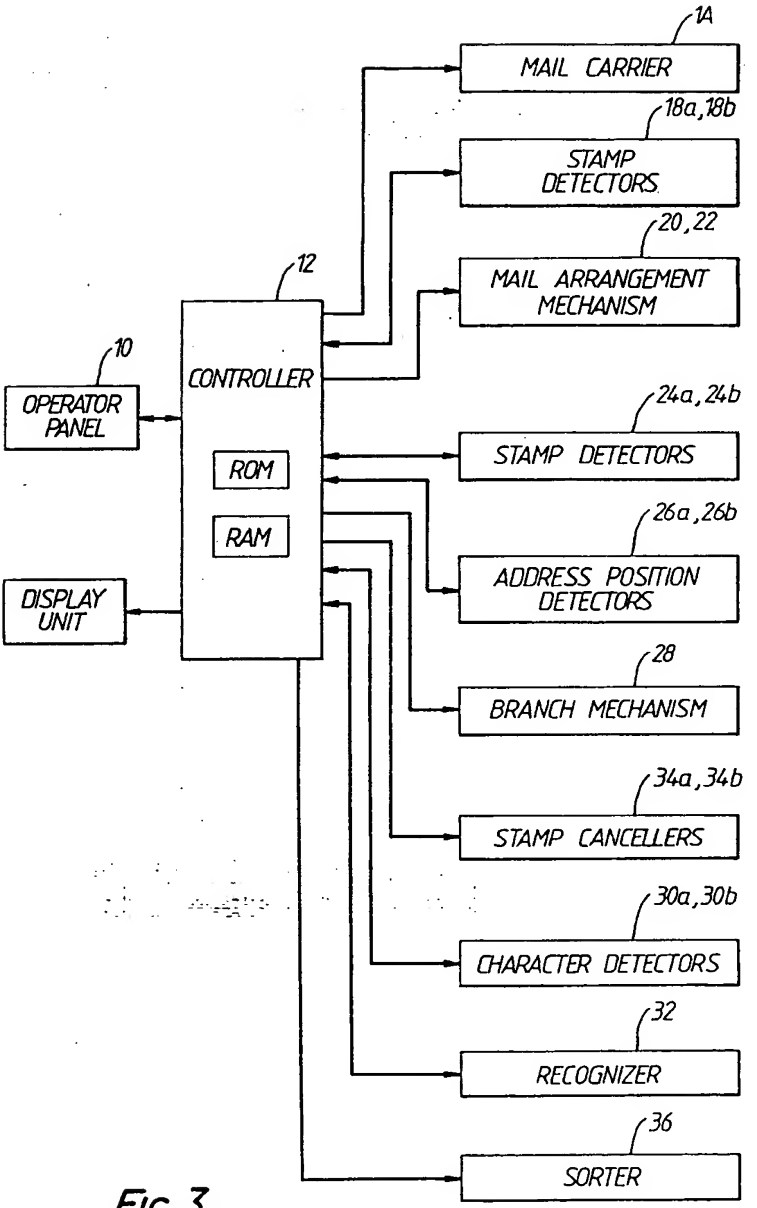

FIG.3.

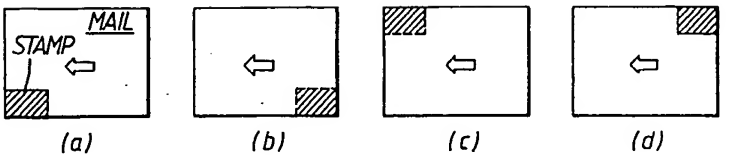

FIG. 4.

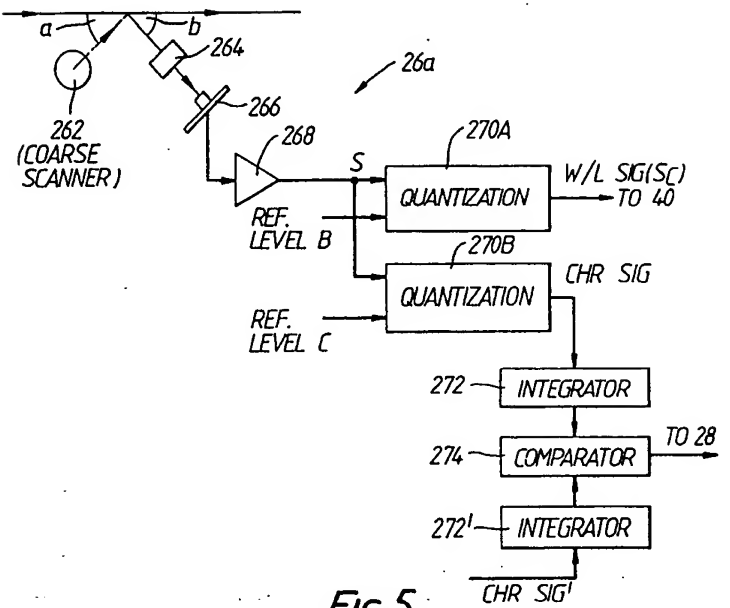

FIG. 5.

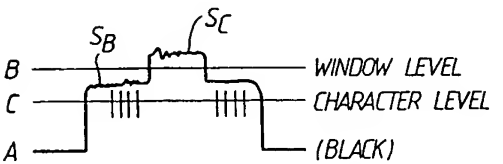

FIG. 6.

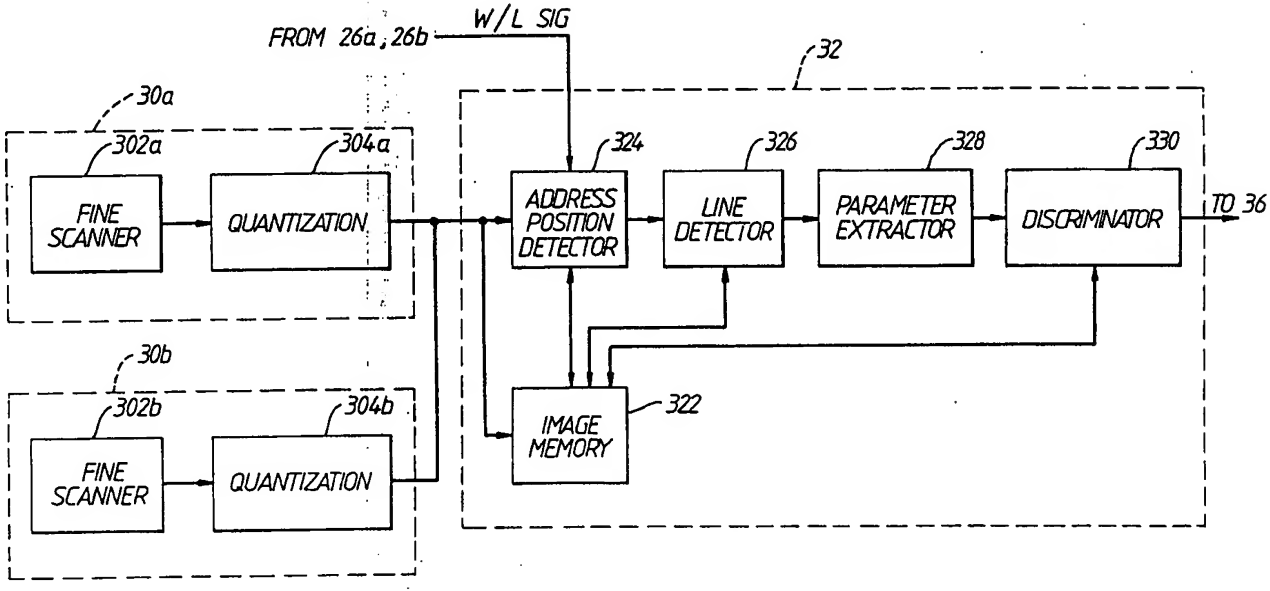

FIG. 7.

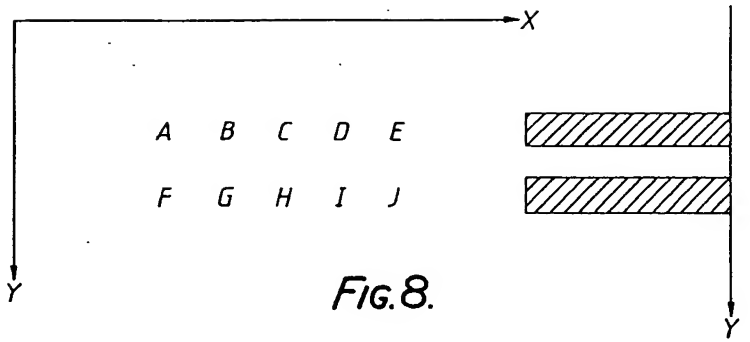

FIG. 8.

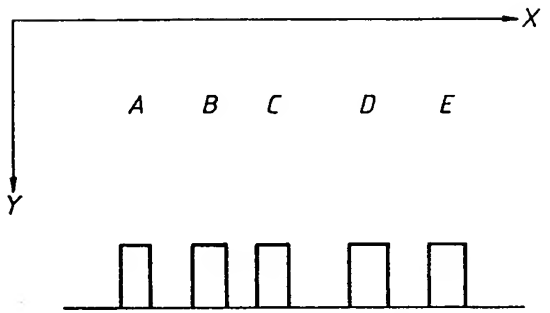

FIG. 9.

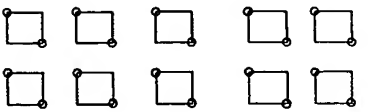

FIG. 10.

0298769

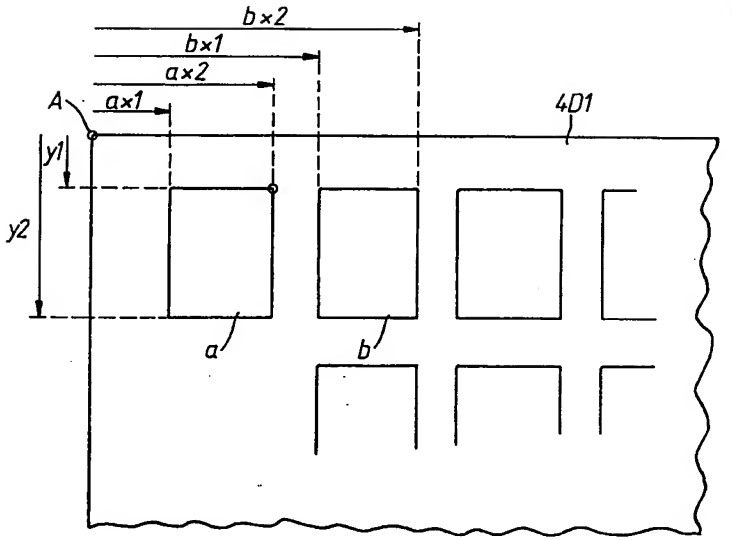

FIG. 11.

FIG. 12.

FIG. 13.

FIG. 14.

FIG. 15.

FIG. 16.

0298769

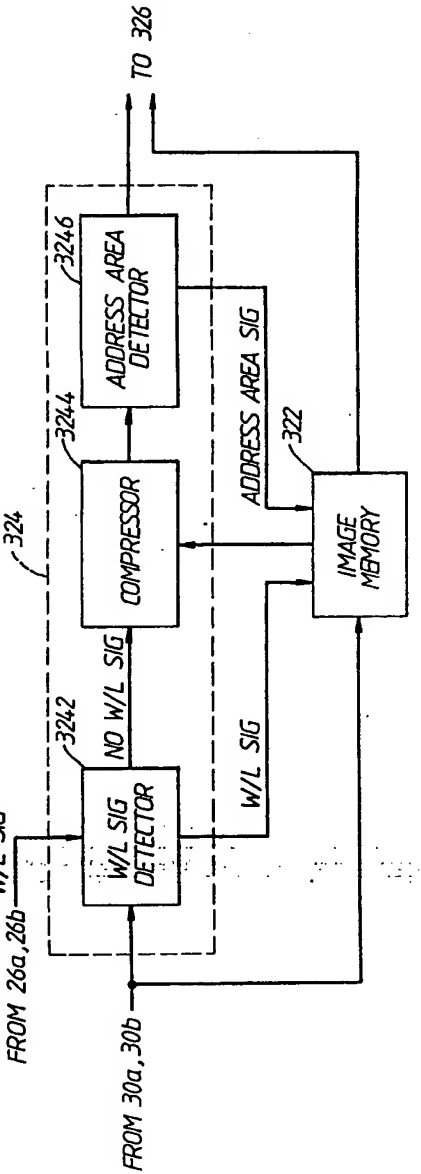

FIG.17.